

April 17, 2014

Volume 3, Issue 4



Image by Espacio telefonica (cc)

## The Turn to Data Analytics and International Law

[Fleur Johns](#)

University of New South Wales

I took a flight the other day. A friend booked and paid cash for it, returning a favour. She was going to join me, but changed plans at the last minute. I requested a special meal. As I settled into my seat, I became associated with my fellow passengers through the sharing of physical space, the exchange of words, the parallel watching of movies, the conjunction of travel plans. We also became actually or potentially co-placed in a pattern.

### *Patterning*

Among the patterns in or against which my co-passengers and I will have been arranged are those configured as lists. Some of the lists in question will have been generated on the global plane (UN Security Council lists identifying those subject to travel bans, for instance). Others will have originated and been maintained nationally (the US government's no-fly list, automatic selectee list and terrorist watch-list, for example, and their global counterparts). In many instances, private data-mining companies will contribute to the bundle (as in many governments' recourse to WorldCheck's database of 'heightened-risk individuals and businesses').

These lists govern conduct and decision-making in their own right, but they are also frequently conjoined with algorithms: encoded procedures or sets of mathematical rules for the processing of data, often with in-built capacity to modify their processing operations on the basis of newly acquired information. Around the world, 'smart border' technologies algorithmically trawl vast, shifting datasets of recorded fragments. Traces corresponding to coded criteria – names, dates of birth and gender, biometric data, certain features in travel itineraries, meal choices, booking and payment practices and the like – are pooled and arranged into clusters. Location within these clusters will be determined by strength of association between particular configurations of data and information gleaned from past events. Precisely the sorts of traces mentioned above – payment with cash, the splitting of a passenger name record (PNR), certain meal

choices – may, when combined with other factors, bring someone to attention within these patterns.

The datasets so combined express a lumpy array of regulatory rationales. A recent article in *Government Security News* observed that the US Office of Biometric Identity Management (OBIM) program – formerly US-VISIT, the United States Visitor and Immigration Status Indicator Technology – identifies ‘terrorists, wanted criminals, sex offenders, immigration violators and international criminals at airports and ports of entry around the world’.<sup>1</sup> The WorldCheck database, mentioned above, classifies individuals and businesses in over a dozen risk categories. The spectrum of potential wrongdoing in relation to which my co-passengers or co-patterners and I may be scrutinized is, accordingly, highly elastic and defies any cohering logic. It is the convergence of ancient and modern analytical technique – the list and the algorithm – that brings these concerns together.

Amid all this sifting and gleaning, any convergence of data generating strong correlations with someone or something flagged somewhere as worthy of ‘interest’ may yield ‘actionable insights’ for immigration and security personnel. People may be channeled into distinct processing routes and subjected to differing levels of scrutiny on this basis. Some of my fellow co-passenger or co-patterners may have paid an annual fee for pre-screening, or secured expedited clearance by other means. The rest of us will have to await allocation to whichever queue or interview process we may be algorithmically assigned, most likely without ever being made aware of the predicates of that assignment.

### ***The List-Plus-Algorithm***

Legal scholars have written a fair amount about the turn to data analytics for global governance from the perspectives of privacy, data protection, international humanitarian law, administrative law, constitutional law and human rights.<sup>2</sup> Relatively little has been said, however, about the juridical forms propagated across the international legal field in this connection: prominent among them, the list-plus-algorithm.

Not only does the list-plus-algorithm recur in the immigration and border security contexts evoked above. It is recurrent too in international environmental law. Listing of animal and plant species in Appendix I or II to the Convention on the International Trade in Endangered Species of Wild Fauna and Flora (CITES), affording differing levels of treaty protection, is the outcome of a process of decision taken by the Conference of Parties to the Convention every two years, with input from State representatives, expert committees and the CITES Secretariat. Yet the unspecified ‘qualitative and quantitative information’ called for in a listing or de-listing proposal will typically include outputs of

---

<sup>1</sup> J. Goodwin, ‘Accenture to Support Data-Sharing and Web Access at DHS office of biometric identity management’ *Government Security News* 10 September 2013, <http://www.gsnmagazine.com/print/33042>.

<sup>2</sup> For some recent examples, see T. Z. Zarsky, ‘Transparent Predictions’ (2013) 4 *University of Illinois Law Review* 1503-1569; K. Crawford & J. Schultz, ‘Big Data and Due Process: Towards a Framework to Redress Predictive Privacy Harms’ (2014) 55 *Boston College Law Review* 93-128.

species distribution modeling (SDM) using software implementing one among a number of possible presence/absence algorithms. In this sense, CITES listing both triggers and is informed by SDM analysis.

A further illustration of this conjunction is the UN's list of individuals or entities associated with Al-Qaida, Usama Bin-Laden and/or the Taliban and subject to sanctions on that basis: a device that originated in UN Security Council Resolution 1267 of October 1999 (the '1267 List'). When nominating individuals or entities for inclusion on this list, UN Member States must provide some 'indicat[ion]' of 'the nature of the association' between the individual or entity in question and Al-Qaida, Usama Bin-Laden and/or the Taliban, including 'specific information' and advice as to 'the nature of the information, for example, intelligence, law enforcement, judicial, media, and admissions by subject'.<sup>3</sup> It is in this context that information gleaned from the algorithmic analysis of open source, unstructured data (including by automatic Web crawlers or 'bots') frequently comes into play, alongside other data-gathering and forecasting techniques.<sup>4</sup> Again, this typically involves both public and private actors and resources: government departments and databases operate in combination with commercial web intelligence software and patented search algorithms, deploying a range of standards for the elucidation of 'associations'.

### ***Lines of Sight***

It is a routine preoccupation of international lawyers that global norms and public decision-making processes should be apparent to those whom they impact: transparency is today treated as a meta-principle of international legal order. The normative architecture, and regulatory ramifications, of the list-plus-algorithm respond unreliably to this imperative. Data moves between accessible and inaccessible forms within and around it, subject to varying rates of turnover, with disparate pathways of input and output, offering variable prospects of inquiry.

Narrative summaries, added to the 1267 List in 2008 in the wake of litigation, purport to offer some window onto analyses underpinning the listing process. Their effect, however, is less revealing than deflective – more bedtime story than reasoned justification. These summaries serve to embolden the secreted knowledge claim with which the list-plus-algorithm comes embedded, bolstering its appeal to trust: 'these are not just some people that turned up in a pattern', the narrative summary insists, 'we *know* these people'.

In CITES, inscrutability is less an incident of secrecy than one of technological complexity and disciplinary division. According to a recent report in *Science*, SDM entails deployment of software the intricacies of which are not grasped by many scientist-modelers, let alone the CITES decision-makers to whom modeling outcomes

---

<sup>3</sup> See <http://www.un.org/sc/committees/1267/listing.shtml>

<sup>4</sup> See, e.g., V.S. Subrahmanian (ed), *Handbook of Computational Approaches to Counterterrorism* (Springer: New York, 2013).

are delivered: there are ‘many in the SDM domain unable to interpret the original algorithms, much less understand how they were implemented in the distributed code’.<sup>5</sup>

In both these settings, the list appears as the public face of a ‘black box’; it gives durable, objective form to processes ongoing and largely inaccessible.<sup>6</sup> Yet, for all its relative straightforwardness, the list remains stubbornly unforthcoming as a knowledge form. As Canadian criminologist Mariana Valverde has observed, lists are typically unsystematic and unprincipled in their make-up; they refuse synthesis.<sup>7</sup> Any insight they afford will, moreover, often be transitory – a snapshot of configurations soon surpassed or likely to be so. In the list-plus-algorithm pairing, the list *seems* the more approachable, but listing and legibility frequently do not align. Lists may make algorithmic outcomes plain, to some degree, but do not explain them.

Data configurations in and around the list-plus-algorithm and passing placements within them are, accordingly, difficult for law and lawyers to see, let alone negotiate. They may, in some instances, be reverse engineered from patterns of conduct. Migration advocates have, for example, sometimes worked up unpublished ‘safe country of origin’ lists from patterns of asylum applicants’ treatment by authorities.<sup>8</sup> Where data governance cannot be rendered navigable by other means, it may be possible to provoke a pattern to the surface through tactical experiment towards the generation of false positives. Rousseau-esque dreams of total transparency will, nevertheless, remain just that in most instances.<sup>9</sup> Lines of sight into decision-making process tend to tangle and dead-end in the proximity of the list-plus-algorithm.

### **Profiling**

There is, nonetheless, much that seems familiar in the encounter with lists-plus-algorithms for global governance. Could this be little more than a new iteration of the actuarially generated profile? If so, countering techniques are on hand, should they be required. International lawyers are accustomed to worrying about discriminatory profiling. When Rosalind Williams Lecraft (a Spanish citizen of African-American descent) was singled out for an identity check by a National Police officer at a Spanish railway station, the Human Rights Committee later opined that she had suffered discrimination contrary to Article 26 of the International Covenant on Civil and Political

---

<sup>5</sup> L. Joppa, G. McInerney, R. Harper, L. Salido, K. Takeda, K. O’Hara, D. Gavaghan, S. Emmott, ‘Troubling Trends in Scientific Software Use’ (2013) 340 *Science* 814-815.

<sup>6</sup> B. Latour, *Science in Action: How to Follow Scientists and Engineers Through Society* (Harvard University Press: Cambridge MA, 1987) 131.

<sup>7</sup> M. Valverde, *Law’s Dream of a Common Knowledge* (Princeton University Press: Princeton, 2003), 159-163, 173-179.

<sup>8</sup> H. Mårtensson & J. McCarthy, “‘In General, No Serious Risk of Persecution’: Safe Country of Origin Practices in Nine European States’ (1998) 11 *Journal of Refugee Studies* 304-325.

<sup>9</sup> J. Starobinski, *Jean-Jacques Rousseau: Transparency and Obstruction*, trans. A. Goldhammer (University of Chicago Press, Chicago, 1988).

Rights.<sup>10</sup> Rights, reporting mechanisms, and random sampling all embody potential correctives for profiling.<sup>11</sup>

The idea of the profile does not, however, seem to capture the operation of the list-plus-algorithm for governance, or risks associated with that operation. Data analytics do not demand or generate a stable profile. Rather, the ‘centroid’ or ‘seed point’ to which associations are drawn in cluster analysis, or the ‘training set’ upon which machine learning algorithms build, will be subject to continual, automated optimization in the face of incoming data. Security lists may specify names of people worthy of suspicion, but these suspicions are not actualized, for the most part, by finding people who look like those named physically, or share their racial or ethnic traits. Rather, what is sought is an intensification or shimmer in the data – a momentary confluence of temporal, spatial, human and financial data points – that bespeaks some mathematical association to past action and thereby projects a possibility of future action. Species lists may enumerate animals and plants comparably proximate to extinction, but this does not evidence likeness in any other sense. Indeed, political science literature on particular CITES listing controversies suggests that the conditions of species listed may be decidedly unlike, but for their having undergone somewhat similar processes of analysis-towards-death.<sup>12</sup>

### **Resizing**

Perhaps regulatory risks that may be associated with lists-plus-algorithms for global governance are best grasped in terms of under- or over-inclusiveness, rather than as profiling? Lawyers everywhere are attuned to these concerns and adept at resizing the reach of rules to address them. There are things that lawyers can readily envisage doing to correct a list-plus-algorithm’s under- or over-inclusiveness: clarify the design brief; add to or subtract from the list or training set; insert some capacity for human complaint and review; or otherwise tinker to try to tighten the correlation between dataset and rationale.

The difficulty with such thinking, in relation to the list-plus-algorithm, is that it presumes that capacity for line-drawing in algorithmic design resides with those to whom over- or under-inclusiveness will become apparent and a matter of concern. Conversely, it assumes that capacity for impact-assessment and the imperative of ongoing justification attach to those who possess relevant line-drawing capacity. Such alignment in capacities seems improbable, given the technical specializations involved. As British geographer Louise Amoore has shown (drawing on interviews conducted with security software designers), whether or not to add or delete an element in an algorithm may depend as much on whether it looks ‘pretty’ on the screen (to those attuned to the

---

<sup>10</sup>Human Rights Committee, Communication No. 1493/2006, *Lecraft v Spain*, CCPR/C/D/1493/2006, 17 August 2009.

<sup>11</sup> B.E. Harcourt, *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age* (University of Chicago Press: Chicago, 2007).

<sup>12</sup> T. Gehring & E. Ruffing, ‘When Arguments Prevail Over Power: The CITES Procedure for the Listing of Endangered Species’ (2008) 8 *Global Environmental Politics* 123-148.

aesthetics of code) as on normative considerations of fit between analytics and rationale.<sup>13</sup> From the perspective of a software designer, contracted to deliver on a brief or perhaps only some portion of a brief, it may not be all that consequential whether the software in question is designed to recommend products to a consumer, or to re-route airline passengers for intensive screening. All may be reduced to service delivery optimization, and often is so reduced in industry parlance. Flow-on considerations of impact are likely to be left to 'end users' who may or may not have a good grasp of the complex and dynamic relationship between code, data and law.

### ***Co-Patterning as Association on a Global Scale***

Lists-plus-algorithms emplace new alignments of people, places and things – or fragmented approximations of the same – on the global plane and suggest new ways of eliciting and conditioning associations among them. Beyond the illustrations given, many matters of international legal import now entail recourse to data analytics including, often, some version of the list-plus-algorithm: the distribution of public welfare and international humanitarian aid; the deployment of military and police forces; assessments of wellbeing, vulnerability, political change and public opinion in particular areas and global policymaking on the strength of these. These measures actualize a set of juridical relations that may prove just as significant as those among fellow citizens, right-holders, or consumers and are not wholly reducible to any of the latter: relations of co-placement in one or other governance pattern.

Talk of terrorism, extinction and automated screening may suggest otherwise, but all is not necessarily worrisome in these developments. The use of data analytics for governance may ensure greater responsiveness and resource allocation to areas of greatest need. Lists facilitate national implementation of global norms, which may (depending on the norms in question) be a good thing. Yet, as with any legal measure, lists-plus-algorithms can generate perverse outcomes, cause unintended consequences, create blindspots, condition for passivity, and legitimize domination, and not only through their misuse. International lawyers have an established repertoire for addressing worries of this kind. Yet this repertoire may not be up to the task of negotiating global relations conducted through the list-plus-algorithm medium, as the foregoing discussion has begun to show.

Perhaps a starting point for expanding our repertoire in this respect could reside in the scene with which I opened: that planeload full of co-passengers or co-patterners. We are accustomed to thinking of ourselves, in such settings, in particular associative formations, conceived for the most part transactionally and on a vertical axis: as individual consumers of travel services purchased from one or other corporation; as individual occupiers of state airspace and crossers of state borders; as exercisers of a qualified freedom of movement opposable to governments. Yet, occasionally, something happens – something humorous or tragic – that occasions the enlivenment of relations otherwise, horizontally, among those contingently brought together in time,

---

<sup>13</sup> L. Amoore, *The Politics of Possibility: Risk and Security Beyond Probability* (Duke University Press: Durham, 2013), 129.

space and analytical pattern. Perhaps in these momentary alignments we catch a glimpse of juridical formations on a global plane with which we are, as yet, only beginning to experiment, and the lawful import of which we are still yet to fully grasp. Global co-patterners we are, but we are also yet to become.

The International Covenant on Economic,  
Social and Cultural Rights

Ben Saul, David Kinley, and Jaqueline Mowbray

**30%**  
Discount\*

\*For ESIL members  
Use code **ALAUTH14**  
when you order at

[www.oup.com](http://www.oup.com)